Detection of Mosquito Breeding Areas using Semantic Segmentation

Pravina Mylvaganam[1], and Maheshi B. Dissanayake[2]

[1,2]*Department of Electrical and Electronic Engineering, University of Peradeniya, Peradeniya, Sri Lanka*

[1] *pravina.m@eng.pdn.ac.lk*, [2] *maheshid@eng.pdn.ac.lk*

Abstract— **The combination of deep learning (DL) and convolutional neural networks (CNN) with image analysis to locate stagnant water will play a crucial role in the dengue control process. This paper aims to automatically segment stagnant water areas in aerial images, acquired by a drone camera, using the latest CNN semantic segmentation method (SegNet). To enhance the effectiveness of our system and as the solution for the lack of dataset, we utilise two different datasets with high domain feature correlation. In our project, pre-training is first done on a large generalised dataset with areas of water, and then the trained model with trained weights is retrained using a task-specific dataset. It should be noted that the conditional distribution of the labels is similar for both datasets. The performance of the SegNet was evaluated with pixel accuracy and dice score. The model exhibited an accuracy of 80% and a dice score of 91%, indicating that our proposed method is efficient to segment water in RGB aerial imagery.**

Keywords—Semantic Segmentation, aerial images, water retaining objects

I. Introduction

Epidemics of dengue fever have been recognized for more than 200 years [1]. One modeling estimate indicates 390 million dengue virus infections per year, of which 96 million (67–136 million) manifest clinically [2]. Another study on the prevalence of dengue discloses that 3.9 billion people are at risk of infection with dengue viruses. Despite a risk of infection in 129 countries, 70% of the burden is in Asia [2]. The principal mosquito vector of dengue and urban yellow fever is Aedes aegypti, which breeds in water that has been collected in natural and artificial containers around human habitations. Water storage tanks, flower pots, garden fountains, bird baths, fridge trays, water dispenser trays, broken cisterns, discarded bottles and tires, tins, coconut shells, etc are all possible sites for mosquitoes to breed. The eggs can survive up to one year in dry containers and hatch when water is available. Therefore, keeping neighborhoods clean and free of receptacles that attract dengue-carrying mosquitos is vital. In addition to vector control measures, World Health Organization is working closely with the Ministry of Health Nutrition and Indigenous Medicine to control the spread of dengue, by specifically collaborating in reviewing dengue control and prevention activities at district and national levels. Even though they have taken plenty of dengue control activities, some limitations also include identifying mosquito breeding sites in inaccessible places, like rooftops and overhead water tanks. We proposed a system for identifying possible breeding places with UAV-based aerial images to cope with this limitation.

An Unmanned Aerial Vehicle (UAV), commonly known as a drone, is a type of aircraft that operates without a human pilot onboard. Recent technologies have allowed for the development of many different kinds of advanced unmanned aerial vehicles used for various purposes. As control technologies improved and costs fell, their use expanded to many non-military applications. These include forest fire monitoring, aerial photography, product deliveries, agriculture, policing and surveillance, infrastructure inspections, and science.

Image segmentation is an essential component in many visual understanding systems. Segmentation plays a central role in a broad range of applications, including medical image analysis, autonomous vehicles, video surveillance, and augmented reality to count a few. Instance segmentation extends the semantic segmentation scope further by detecting and delineating each object of interest in the image (e.g., partitioning of individual persons). Compared to other techniques such as object detection in which no exact shape of the object is known, segmentation exhibits pixel-level classification output providing richer information, including the object's shape and boundary. A recently emerged semantic segmentation method, which incorporates CNN structure is SegNet. Semantic segmentation makes it easier to understand images because it segments images into semantically significant objects and assigns each part of predefined labels. Thereby, different objects from remotely captured images can be extracted simultaneously. SegNet method is applied in several remote sensing applications [4]. (Du et al., 2018) [4] exploited the SegNet technique to classify and extract cropland in high-resolution remote sensing images, showing that the proposed approach efficiently obtained accurate results (98%) for the segmentation task.

Although several previous works are available in the public domain for identifying water pooling sites [5], they were designed to address a different task than ours. Mettes et al. developed a robust water detection algorithm for videos. They have proposed a methodology to detect water using spatial and temporal dynamics of water [5]. However, this approach only discusses the identification of water, which spreads in a considerably large area, such as pools and ponds, but does not analysis the applicability of this technique for small-scale water pooling areas.

Our research object is to detect possible water retention areas using aerial images. To find a solution to our problem, first, we have created a dataset using locally collected high-resolution images taken from a UAV operated at low altitudes. These characteristics of the dataset bring more clarity to the task as well as novelty. Furthermore, it helps deep learning models to learn specific features which assist to make accurate and precise detection of even small water retention areas.

With this objective, this paper aims to automatically segment stagnant water areas in RGB aerial imagery using the SegNet semantic segmentation method. We experimented with our custom dataset captured by a drone camera. Since our task-specific dataset is of a smaller size, we used domain adaptation for this task to enhance the efficacy of our system. In our proposed solution, pre-training of SegNet is first carried out on a generalized large custom dataset. Later, the trained model is sub-trained on another dedicated dataset collected locally without retraining from scratch.

The rest of the paper is organized as follows. Section 2 and Section 3 presents the methodology adopted in this study and discusses the results obtained in the experimental analysis respectively. Finally, Section 4 summarizes the main conclusions.

II. Methodology

In this paper, we proposed a semantic segmentation model for identifying mosquito breeding sites that contain water efficiently and automatically. At the initial stage of the research, two different datasets were acquired using images captured with a drone camera during a clear sunny day. The model is generalised to all geographical locations with the image capturing condition set as sunny day, as the training was done using universal dataset. For the pre-training purpose, a total of 2,668 images, belonging to one class: water areas, were collected and a total of 300 images of water retaining objects, such as water storage tanks, flower pots, garden fountains, bird baths, water dispenser trays, broken cisterns, discarded bottles and tires, tins, and coconut shells, were captured as a second dataset. Even though the image resolution is 1080 x 1900 pixels for the image samples of the dataset, we evaluated the input images at 256 x 256 during the experiments. Both datasets have been annotated using the LabelMe Software. The third and second column in Fig. 3 shows examples of original and labeled images of the datasets respectively. Next, the first dataset was divided into a training set and a validation set with a 70:30 ratio respectively, while the second dataset was split with an 80:20 ratio respectively.

A. Semantic Segmentation

In this application, the CNN-based SegNet architecture was first trained to segment the pixels into water area and background on a large generalized dataset. SegNet consists of a symmetrical encoder-decoder followed by a pixel-wise classifier as shown in Fig. 1 [3]. SegNet has 13 convolutional layers in the encoder/contraction and the decoder/expansion part, and the designed network consists of 29,443,142 trainable parameters. We chose the hyperparameters to obtain high segmentation accuracy under a low learning rate. The model was trained over 100 epochs with batch sizes of 8 and an adadelta optimizer. Once the model is trained using the large dataset, the trained weight files are migrated to the encoder, which is sub-trained using the smaller yet customized (specific) dataset. This approach is known as transfer learning-based domain adaptation in deep learning.

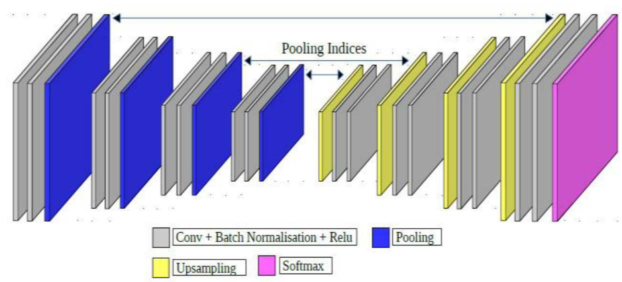


Fig. 1 Proposed SegNet architecture

The same aforementioned CNN SegNet was used for the sub-training with the second dataset. The validation dataset was used to determine the learning rate, which defines how the weights are adjusted in the CNN during training to reduce the risk of overfitting. The model was trained over 100 epochs with the softmax classifier and Google Collab with GPU was used for the simulations. Finally, the test images captured using a drone camera were tested using the trained model to evaluate the model performance.

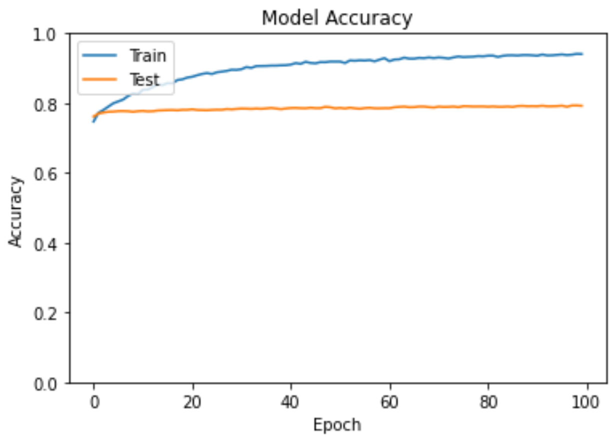
III. Results and Discussion

The performance of our proposed SegNet model is evaluated on RGB test images captured by a drone camera. When we input RGB aerial images, the trained network showed a pixel accuracy of 80% and a training accuracy of 94%. The pixel accuracy shows the percentage of the pixels that were correctly classified. Training time and predicted time were 20 minutes and 8 sec respectively. Furthermore, model accuracy and loss variation through the 100 epochs during the final training is shown in Fig. 2. In addition, Fig. 3 shows samples of RGB test images and segmented outputs of the SegNet model.

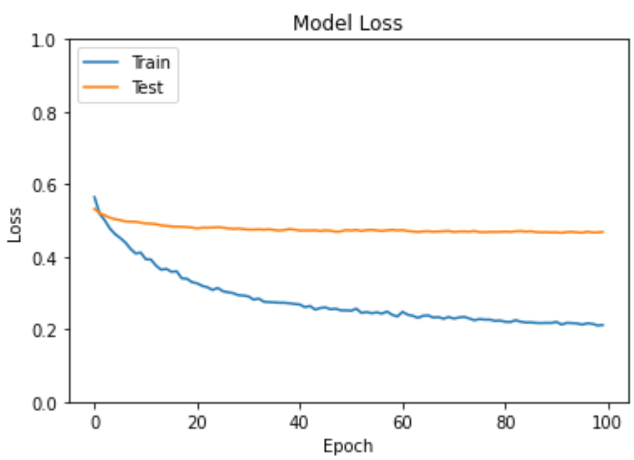
Although the loss variation shown in Fig 2 (b), shows room for further improvement, the sample test outputs in Fig. 3 are aligned with the expected results. Furthermore, accuracy pixel estimates indicate that the method used is efficient to segment water in drone camera images with a small amount of dataset with the help of transfer learning approaches.

Furthermore, the Dice score, mathematically expressed as shown in Eq (1), is used to quantify the performance of the proposed image segmentation method. The average dice score of 91% indicates that the pixel-wise degree of similarity between the model predicted segmentation mask and the ground truth is high.

$$\text{Dice coefficient} = \frac{2*True\ positive}{2*True\ positive+False\ positive+False\ negative} \quad (1)$$



(a)



(b)

Fig. 2 (a)Training and validation accuracy variation of SegNet model; and (b) Training and validation loss of SegNet model

IV. CONCLUSION

The main objective of this research is to propose a mechanism to identify possible water retention areas in inaccessible places. Especially on rooftops. With this objective, we proposed a system to automatically detect potential Aedesaegypti breeding sites using a Deep Learning algorithm to help public health agents to combat its reproduction. We present a semantic segmentation method (SegNet) to automatically segment water in imagery acquired by a drone camera to identify the mosquito breeding sites. The overall performance, with an average accuracy of 80% and average dice score of 91%, indicated that the SegNet method is an efficient approach to segment the water area in images.

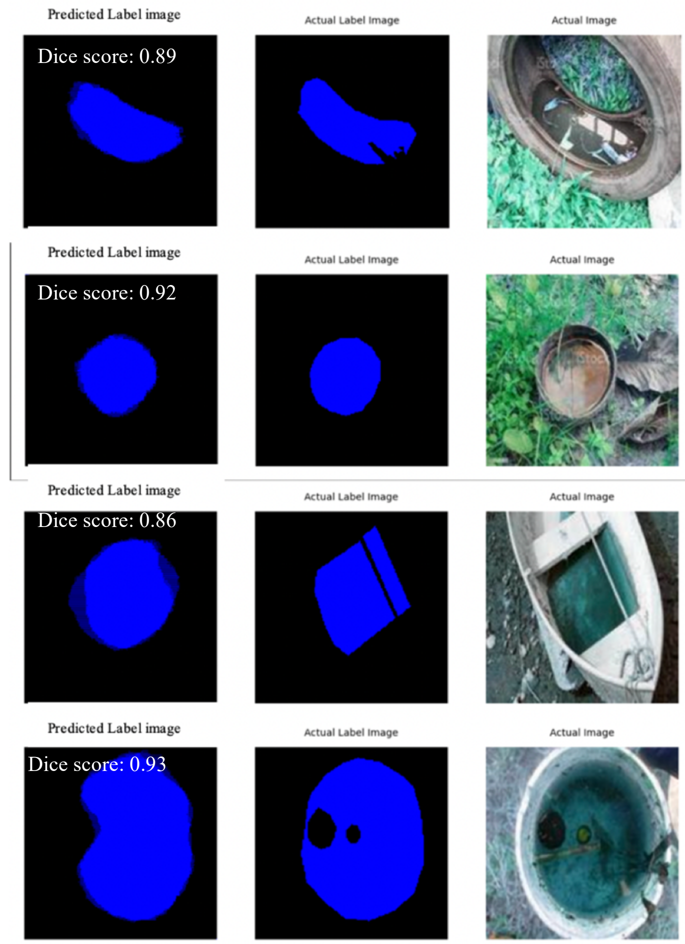


Fig. 3 Sample original RGB images; Actual labeled images; and Segmented outputs from the SegNet model

REFERENCES

[1] Siler JF, Hall MW, Hitchens AP, "Dengue: its history epidemiology, mechanisms of transmission, etiology, clinical manifestations, immunity, and prevention," *Philippine J Sci. 1926*, vol. 29, no. 1-2, pp. 1-304.

[2] Bhatt, S., et al., "The global distribution and burden of dengue," *Nature*, vol. 496, no. 7446, pp. 504–507, April 2013.

[3] Badrinarayanan, V., Kendall, A., Cipolla, R., "SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation," *IEEE Trans. on Pattern Anal. and Mach. Intel.,* vol. 39, no. 12, pp. 2481–2495, 2017.

[4] Du, Z., Yang, J., Huang, W., Ou, C., "Training SegNet for cropland classification of high-resolution remote sensing images," in *Proc. AGILE Conf.*, 2018.

[5] P. Mettes, R. T. Tan, and R. C. Veltkamp, "Water detection through spatio-temporal invariant descriptors," *Comput. Vis. Image Underst.*, vol. 154, pp. 182–191, Jan. 2017.